# The Machine Learning Revolution in AI

Luc De Raedt
*Dept. Computer Science*
*luc.deraedt@cs.kuleuven.be*

KATHOLIEKE UNIVERSITEIT LEUVEN

erc

***Observation:***

*Machine Learning is everywhere in AI and Computer Science*

# Overview

What is machine learning ?

Why is it useful ?

What is behind it ?

# *What is Machine Learning ?*

# Machine Learning

A *machine* learns when

it improves it *performance*

on a specific *task*

with *experience*

*Central to Artificial Intelligence & Cognitive Systems*

# AlphaGo

# AlphaGo

Machine = AlphaGo Computer Program

Task = playing GO

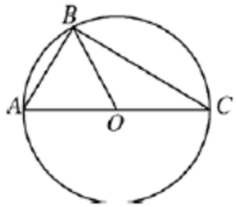
Performance = % of won games

Experience = huge data base of games

# Euclid (Allen AI Institute)

**GeoS Demo – An End to End Geometry Problem Solver**

In the figure above, triangle ABC is inscribed in the circle with center O and diameter AC. If AB=AO, what is the degree measure of angle ABO?




(A) 15°
(B) 30°
(C) 45°
(D) 60°
(E) 90°

Solve Problem

# Euclid

Machine = Euclid Computer Program

Task = answering SAT questions

Performance = % of correct answers

Experience = reading the web, textbooks, past examples

# Watson / Jeopardy


*machine reading*

# Robot Scientist

King et al.  Science 2009

# Robot Scientist

Machine = Robot

Task = find highest activity molecule

Performance = nof. experiments / reliability of model

Experience = past experiments and data

# Automated Data Science ?

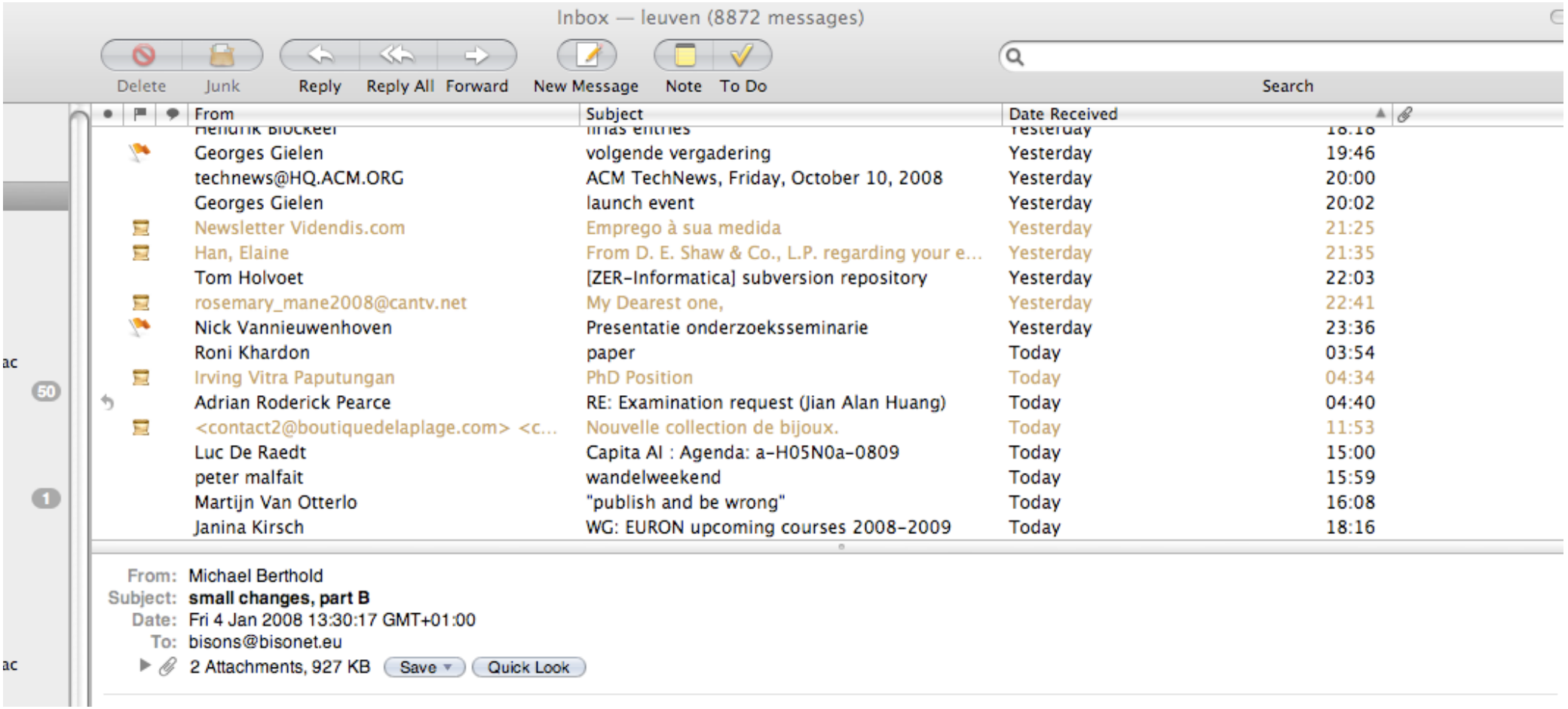
## Can we automate Machine Learning ?

## Can we apply the idea of the robot scientist to Machine Learning and Data Science ?

| Product | State | Sep | Oct | Nov | Sub-total | | Product: | Sales: |
|---|---|---|---|---|---|---|---|---|
| Cherries | CA | 200 | 360 | 230 | 790 | | Apples | 2070 |
| Bananas | AZ | 350 | 230 | 150 | 730 | | Bananas | 2420 |
| Apples | TX | 180 | 270 | 200 | 650 | | Lemons | 740 |
| Bananas | KS | 400 | 240 | 310 | 950 | | | |
| Lemons | AL | 250 | 360 | 130 | 740 | | | |
| Apples | FL | 120 | 120 | 380 | 620 | | | |
| Bananas | LA | 330 | 270 | 140 | 740 | | | |
| Apples | KY | 110 | 320 | 370 | 800 | | | |

- What are the formulas here?
- `T1[:, 6] = SUM(T1[:, 3:5], row)`
- `T2[:, 2] = SUMIF(T1[:, 1]=T2[:, 1], T1[:, 6])`

# Spam Filter


Inbox — leuven (8872 messages)
Delete
Junk
Reply
Reply All
Forward
New Message
Note
To Do
Search
From
Subject
Date Received
Georges Gielen volgende vergadering Yesterday 19:46
technews@HQ.ACM.ORG ACM TechNews, Friday, October 10, 2008 Yesterday 20:00
Georges Gielen launch event Yesterday 20:02
Newsletter Videndis.com Emprego à sua medida Yesterday 21:25
Han, Elaine From D. E. Shaw & Co., L.P. regarding your e... Yesterday 21:35
Tom Holvoet [ZER-Informatica] subversion repository Yesterday 22:03
rosemary_mane2008@cantv.net My Dearest one, Yesterday 22:41
Nick Vannieuwenhoven Presentatie onderzoeksseminarie Yesterday 23:36
Roni Khardon paper Today 03:54
Irving Vitra Paputungan PhD Position Today 04:34
Adrian Roderick Pearce RE: Examination request (Jian Alan Huang) Today 04:40
<contact2@boutiquedelaplage.com> <c... Nouvelle collection de bijoux. Today 11:53
Luc De Raedt Capita AI : Agenda: a-H05N0a-0809 Today 15:00
peter malfait wandelweekend Today 15:59
Martijn Van Otterlo "publish and be wrong" Today 16:08
Janina Kirsch WG: EURON upcoming courses 2008-2009 Today 18:16
From: Michael Berthold
Subject: small changes, part B
Date: Fri 4 Jan 2008 13:30:17 GMT+01:00
To: bisons@bisonet.eu
2 Attachments, 927 KB
Save
Quick Look

# Spam Filter

Machine = e-mail program, spamfilter

Task = classify e-mails

Performance = accuracy

Experience = your past input

*Why is it useful ?*

# Why Machine Learning ?

Artificial Intelligence

Machine Learning is very practical

- some programs are too complex to program by hand (spam, go … )
- adaptation and personalisation
- analysing data (data mining), discovering new knowledge

*"we are drowning in data but starving for knowledge"*

# Numerous applications

The enabling technology in AI

- natural language processing
- computer vision
- robotics
- self-driving cars
- data mining / data analysis
  - medical, financial, biology, chemistry, engineering, analytics, …

# Machine Learning and Computer Science

Research in machine learning is affecting ALL areas of computer science and ICT

- programming
  - languages that learn & programming by optimisation
- human computer interaction - personalization SIRI
- adaptive compilers / configuration / operating systems / chips …

*How does Machine Learning work ?*

# How does it work?

Machine learning is all about learning functions

f(input) => output.

different types of functions

different types of data

different criteria (loss functions)

and there are

different schools of thought in ML

# More formally

**Given**

a space of possible instances *Input*

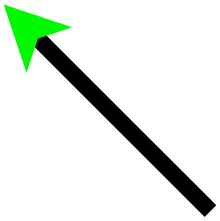
an unknown target function f: *Input* → *Output*

a (hypotheses) space of possible functions L from *Input to Output*

a set of examples E = { (input X, output f(x)) }

a loss function *loss*(h,E) → $\mathbb{R}$

**Find** h ∈ L that minimizes *loss*(h,E)



supervised

*Machine Learning = Variations on this problem*

# Learning functions

Spam

f(e-mail) = spam of geen spam

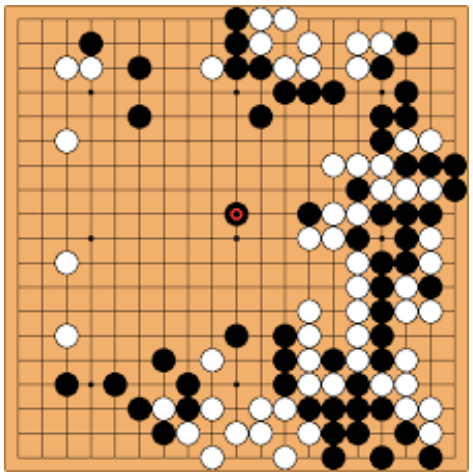
Go

f( board ) = 9-13 predict move (policy)

f( board, 9-13 ) = good /bad

P( board,9-13) = probability of play

Q( board,9-13) = expected "reward"



# What kind of data ?

learning from examples *(supervised / unsupervised)*

good/bad moves ? just moves ?

learning by imitation *(Behavioral cloning)*

imitate de world champion

learning from rewards *(Reinforcement learning)*

just play, reward = board config. / wins / losses

the whole AI problem in a nutshell

*Where does the data come from ?*

# Learning simultaneously

GO :

learn which moves champions can take (classification)

learn by reinforcement by playing

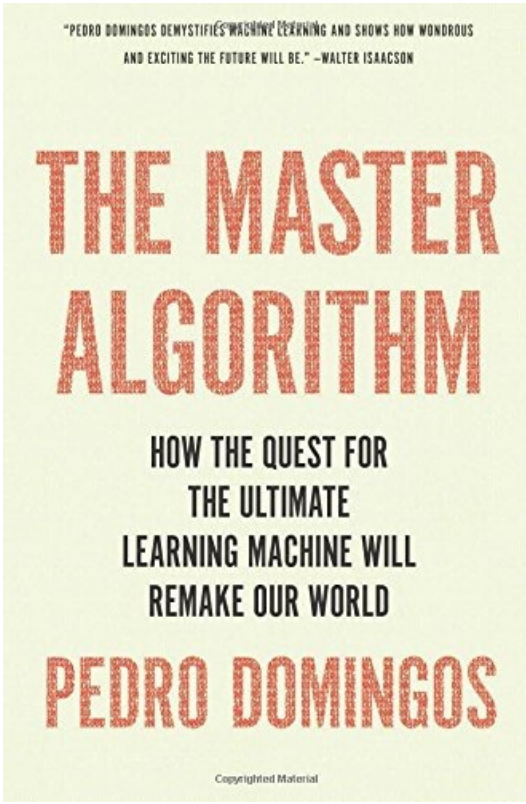
Euclid

learn to map language -> model

learn to solve model

# The five tribes of ML

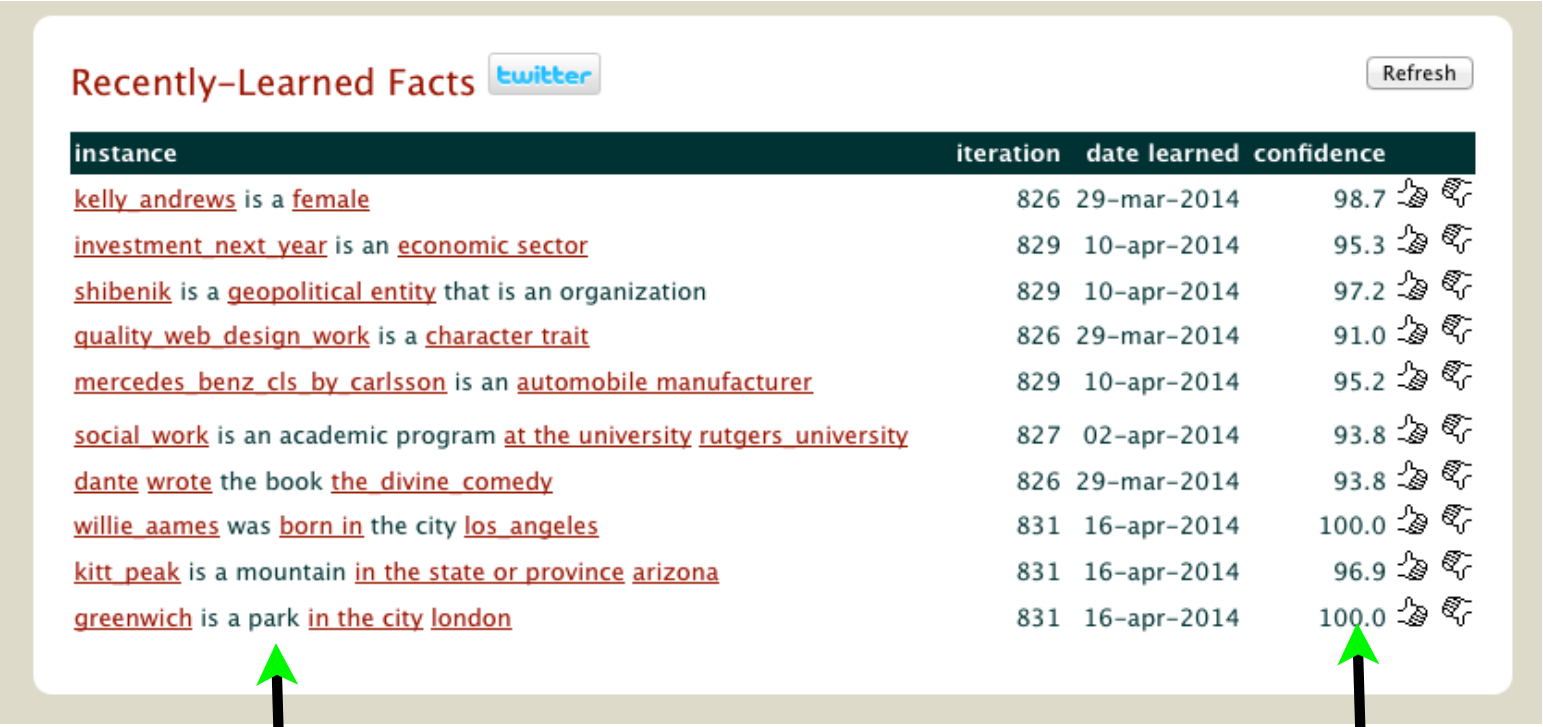
| Tribe | Origins | Master Algorithm |
|---|---|---|
| Symbolists | Logic, philosophy | Inverse deduction |
| Connectionists | Neuroscience | Backpropagation |
| Evolutionaries | Evolutionary biology | Genetic programming |
| Bayesians | Statistics | Probabilistic inference |
| Analogizers | Psychology | Kernel machines |




According to Pedro Domingos

NELL: http://rtw.ml.cmu.edu/rtw/

# Logic (machine reading)



Recently-Learned Facts twitter Refresh

| instance | iteration | date learned | confidence |
|---|---|---|---|
| kelly_andrews is a female | 826 | 29-mar-2014 | 98.7 |
| investment_next_year is an economic sector | 829 | 10-apr-2014 | 95.3 |
| shibenik is a geopolitical entity that is an organization | 829 | 10-apr-2014 | 97.2 |
| quality_web_design_work is a character trait | 826 | 29-mar-2014 | 91.0 |
| mercedes_benz_cls_by_carlsson is an automobile manufacturer | 829 | 10-apr-2014 | 95.2 |
| social_work is an academic program at the university rutgers_university | 827 | 02-apr-2014 | 93.8 |
| dante wrote the book the_divine_comedy | 826 | 29-mar-2014 | 93.8 |
| willie_aames was born in the city los_angeles | 831 | 16-apr-2014 | 100.0 |
| kitt_peak is a mountain in the state or province arizona | 831 | 16-apr-2014 | 96.9 |
| greenwich is a park in the city london | 831 | 16-apr-2014 | 100.0 |

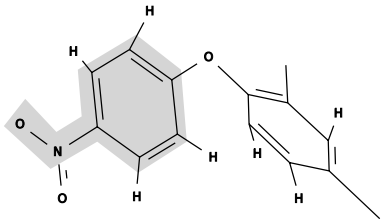
instances for many different relations

degree of certainty

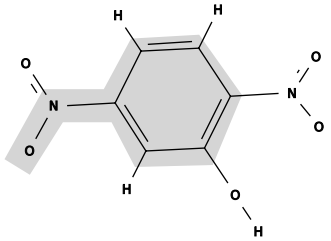
*Learn rules :*

IF Person was Born in City AND City Lies-in Country
THEN Person was Born in Country
IF Person was Born in Country
THEN Person has Nationality Country

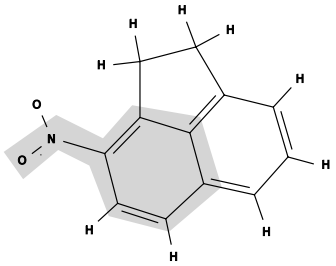
# Structure Activity Relationship Prediction



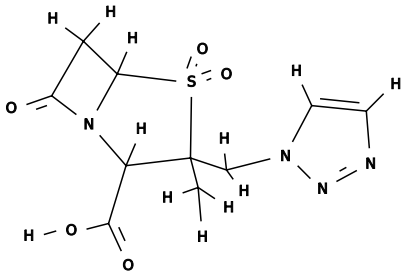
Mutagenic



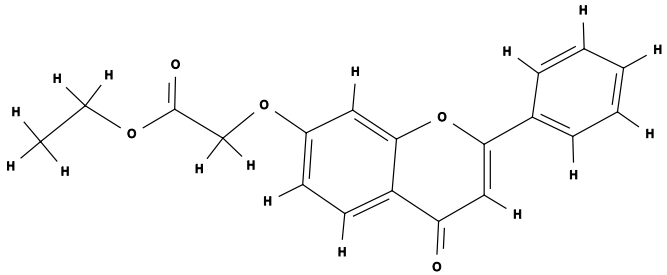
Mutagenic



Mutagenic



Clean



Clean

Molecules are graphs

Patterns are subgraphs

Search for “best”/ most relevant patterns

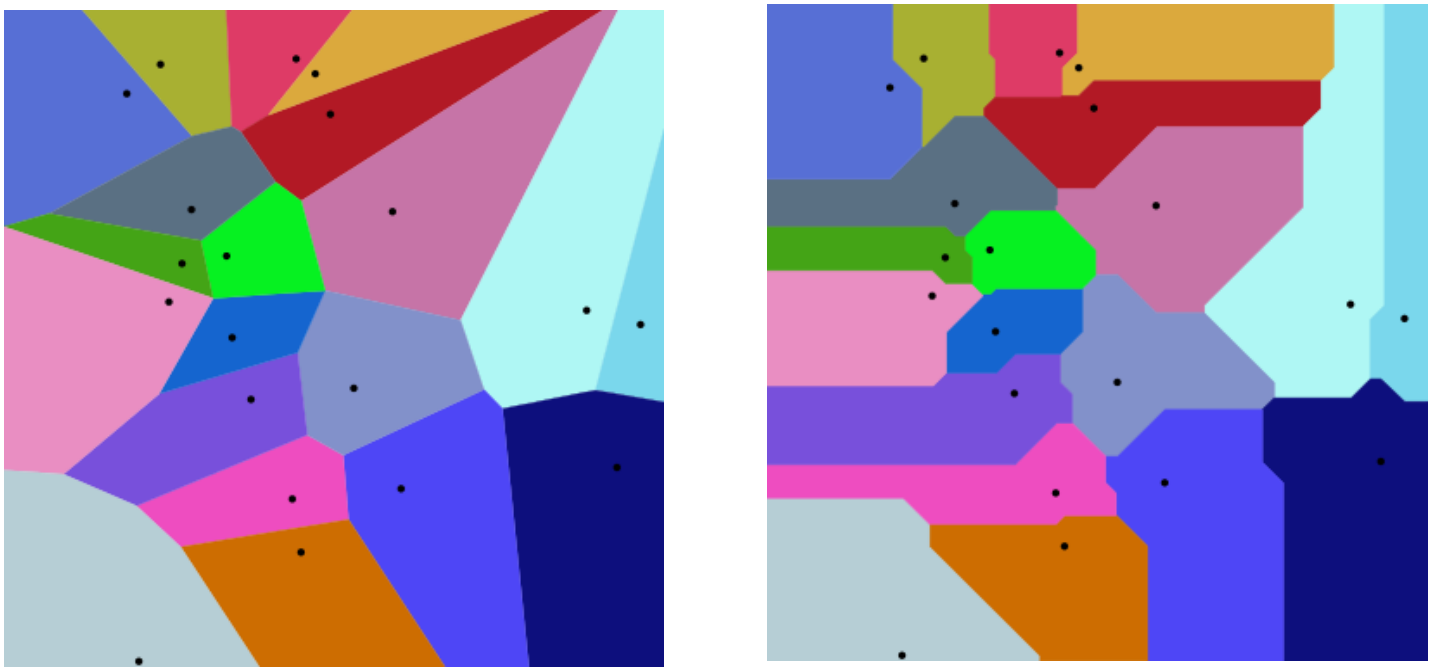
*Figure: Siegfried Nijssen*

Interpretable knowledge

# Analogizers
# Nearest Neighbor

Similarity / Distance measures / Kernels

k- Nearest Neighbor / Support Vector Machines



*Voronoi Diagrms*

# Probabilistic / Bayesian

*Posterior* :
how likely is the function
given the data ?

*Likelihood*:
how likely is it
that the function
generates the data ?

*Prior*
how likely
is this function ?

$$P(Function|Data) = \frac{P(Data|Function) \times P(Function)}{P(Data)}$$

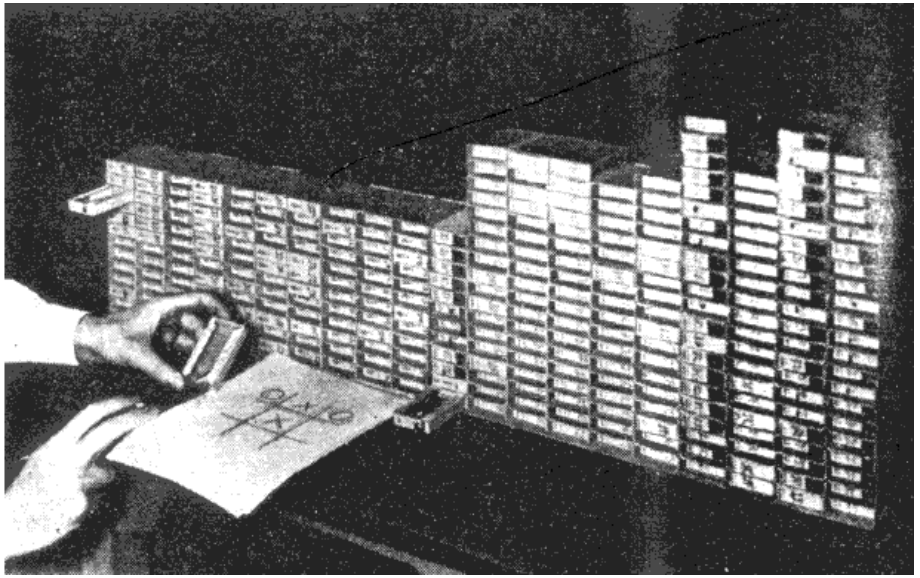
*Law of Reverent Thomas Bayes.*

*Probabilistic graphical models.*

# *An early reinforcement learning system*

# Donald Michie's Menace

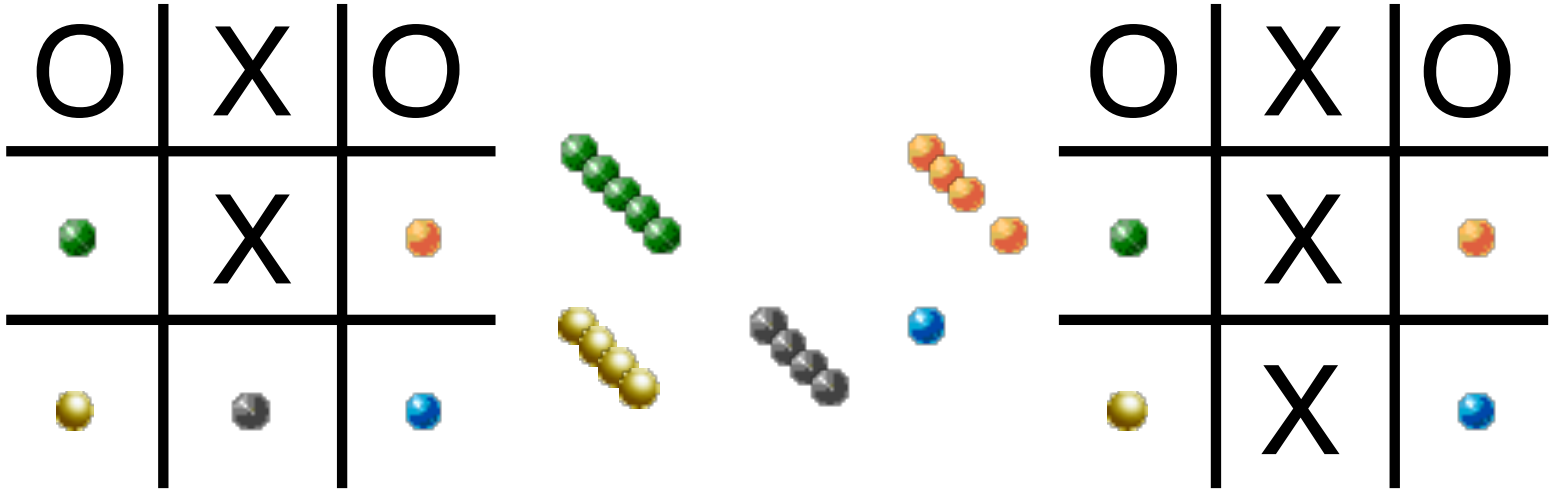
Donald Michie (2007)



Menace (1961)

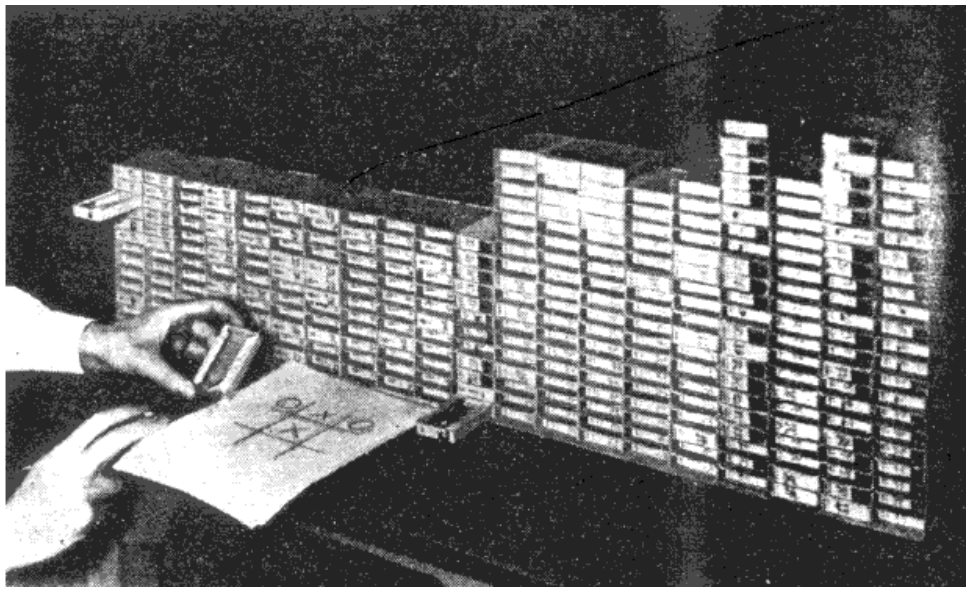
*Machine Educable Noughts And Crosses Engine*

X's move




Choose box on the basis of current position

Execute move



Select pearl from box at random

# Menace

Machine = 287 “boxes” + pearls

Encodes probabilistic function

P(box, color) = probability of move

Learning a function $Q^*(s,a) = R(s,a) + \gamma \sum_{s'} P(s'|s,a) \max_{a'} Q^*(s',a')$

upon *loss*: retain all used pearls

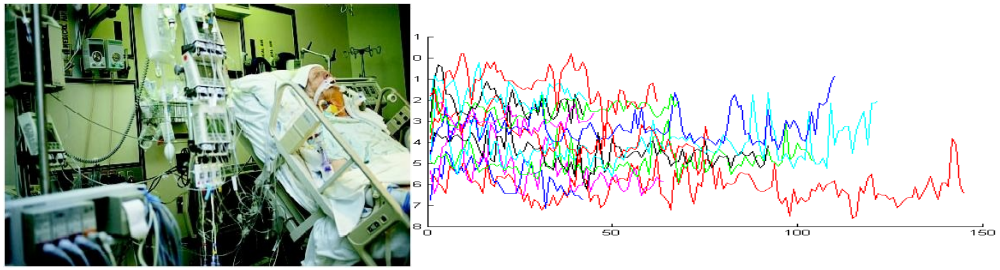
upon *winning*: put used pearls back and an extra one of the same color

Richard Belmann

# Some Applications

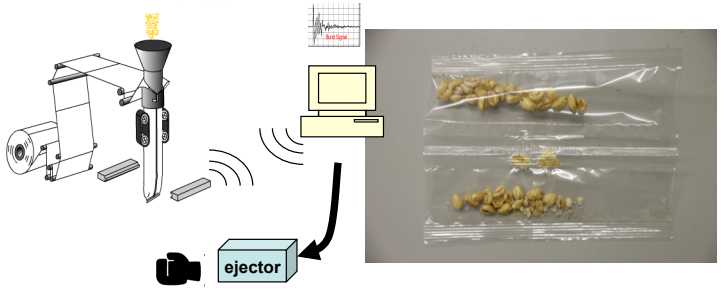
## Intensive Care

with University Hospital



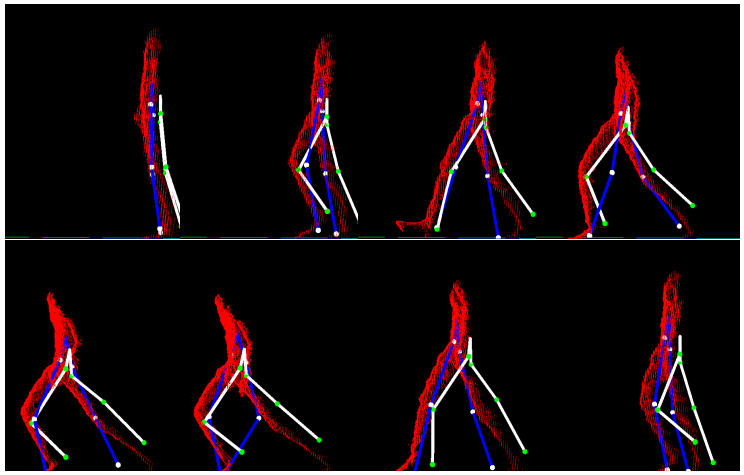
## Predictive Maintenance

with FMTC / Flanders Make / Sirris




## Sports Analytics

Fac. Movement Science and Belgian clubs and leagues

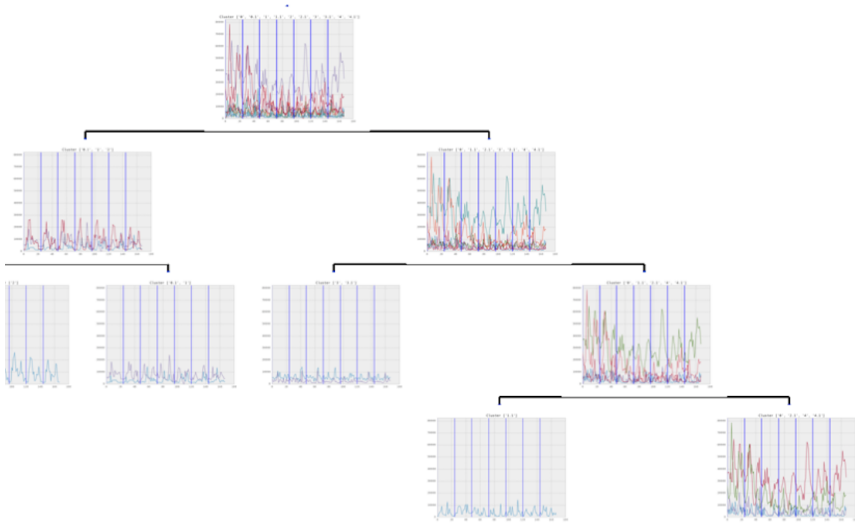


# Energy

in homes

in a shopping chain

in data centers / windmill parks



# ML in Industry

ML and AI are broadly applicable

Numerous success stories — Amazon, Google, IBM, Apple SIRI, Microsoft, self driving cars, Spotify, …

a lot of hype around it, but it is not magic (even though it appears so)

serious applications need serious thought / expertise and investment

strategic decision and vision needed from the start …

US / UK companies are investing …

# Take-away

*Enabling technology for AI, Computer Science and Data Science— this is the era of big data*

*strategy + expertise needed to make successful ML/AI applcations*

*Machine learning is about learning functions from data*

*Different schools in machine learning*



http://elucidata.be/about-elucidata

Contact
wannes.Meert@
cs.kuleuven.be




http://www.sirris.be/hymop